NORMAL FORMS FOR REAL QUADRATIC FORMS

BERNHARD KRÖTZ AND HENRIK SCHLICHTKRULL

ABSTRACT. We investigate the non-diagonal normal forms of a quadratic form on $\mathbb{R}^n$, in particular for $n = 3$. For this case it is shown that the set of normal forms is the closure of a 5-dimensional submanifold in the 6-dimensional Grassmannian of 2-dimensional subspaces of $\mathbb{R}^5$.

Date: April 29, 2008.

1. **Introduction**

According to the principal axes theorem every real quadratic form in n variables allows an orthogonal diagonalization with normal form

$$A_1 x_1^2 + \cdots + A_n x_n^2,$$

where $A_1, \ldots, A_n \in \mathbb{R}$. In this article we investigate (for the case $n = 3$) the existence of other normal forms.

To be more precise, let $q_1, \ldots, q_n$ be quadratic forms on $\mathbb{R}^n$. If for every quadratic form q on n-dimensional Euclidean space there exists an orthonormal basis in which q takes the form

$$q(x) = A_1 q_1(x) + \cdots + A_n q_n(x)$$

for some set of coefficients, we say that this expression is a *normal form* of q.

Passing to matrices, let us consider $V = \mathrm{Sym}(n, \mathbb{R})$, the vector space of symmetric $n \times n$-matrices. On V there is the natural action of the special orthogonal group $K := SO(n, \mathbb{R})$ by conjugation, say

$$k \cdot X := kXk^{-1} \qquad (k \in K, X \in V)\,.$$

If D denotes the space of diagonal matrices in V, then the principal axes theorem asserts that

$$V = K \cdot D := \{k \cdot d \mid k \in K, d \in D\}.$$

Furthermore if $d_1, d_2 \in D$, then $K \cdot d_2 = K \cdot d_1$ if and only if d_2 is obtained from d_1 by a permutation of coordinates (the set of eigenvalues is unique).

The question we address is, *for which n-dimensional subspaces W in V is $V = K \cdot W$?* It would be tempting to assert that the unique property of D (and its conjugates by K), which causes the principal axes theorem, is that it is *abelian*. However this is not correct, in fact there exist non-abelian n-dimensional subspaces W with $V = K \cdot W$ (see the theorem below).

There is some redundancy in the problem, namely the center of V on which K acts trivially. Let us remove that and define $\mathfrak{p} := V_{\mathrm{tr}=0}$ to be the space of zero-trace elements in V. Likewise we set $\mathfrak{a} := D_{\mathrm{tr}=0}$. The principal axes theorem now reads as

$$\mathfrak{p} = K \cdot \mathfrak{a}\,.$$

Theorem 1.1. *Let* $\mathfrak{p} = \mathrm{Sym}(3, \mathbb{R})_{\mathrm{tr}=0}$ *and* $K = \mathrm{SO}(3, \mathbb{R})$. *Define*

$$W := \left\{ X_{\mu\lambda} := \begin{pmatrix} \mu & 0 & \lambda \\ 0 & -\mu & 0 \\ \lambda & 0 & 0 \end{pmatrix} \mid \mu, \lambda \in \mathbb{R} \right\}.$$

Then $K \cdot W = \mathfrak{p}$.

Proof. A more general result will be established later. Here we can give a simple proof.

Let $A \in \mathfrak{p}$ be given, and let $\nu_1 \geq \nu_2 \geq \nu_3$ be its eigenvalues. Then $\nu_1 + \nu_2 + \nu_3 = 0$, and hence $\nu_1 \geq 0 \geq \nu_3$. Let

$$\mu = -\nu_2 = \nu_1 + \nu_3 \quad , \quad \lambda = \sqrt{-\nu_1\nu_3}.$$

The matrix $X_{\mu\lambda}$ has the characteristic polynomial

$$\begin{aligned} \det \begin{pmatrix} \mu - x & 0 & \lambda \\ 0 & -\mu - x & 0 \\ \lambda & 0 & -x \end{pmatrix} &= (-\mu - x)(-(\mu - x)x - \lambda^2) \\ &= (\nu_2 - x)(x - \nu_1)(x - \nu_3) \end{aligned}$$

Hence A and $X_{\mu\lambda}$ have the same eigenvalues, and thus they are conjugate. □

Corollary 1.2. *Every trace free real quadratic form in three variables allows a normal form of the type*

$$A(x^2 - y^2) + Bxz$$

for $A, B \in \mathbb{R}$.

Let us more generally consider a real semi-simple Lie algebra $\mathfrak{g}$ with Cartan decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{p}$. The space $\mathfrak{p} = \mathrm{Sym}(n, \mathbb{R})_{\mathrm{tr}=0}$ is obtained in the special case $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{R})$. Let $\mathfrak{a} \subset \mathfrak{p}$ be a maximal abelian subspace and $K = e^{\mathrm{ad}\,\mathfrak{k}}$. According to standard structure theory of semi-simple Lie algebras the following generalization of the principal axes theorem holds:

- $\mathfrak{p} = K \cdot \mathfrak{a}$.
- $K \cdot X = K \cdot Y$ for $X, Y \in \mathfrak{a}$ if and only if $\mathcal{W} \cdot X = \mathcal{W} \cdot Y$ where $\mathcal{W} = N_K(\mathfrak{a})/Z_K(\mathfrak{a})$ is the Weyl group.

Let $r := \dim \mathfrak{a}$ be the real rank of $\mathfrak{g}$. We consider $\mathrm{Gr}_r(\mathfrak{p})$ the Grassmannian of r-dimensional subspaces in $\mathfrak{p}$. Inside of $\mathrm{Gr}_r(\mathfrak{p})$ we consider the subset

$$\mathcal{X} := \{W \in \mathrm{Gr}_r(\mathfrak{p}) \mid K \cdot W = \mathfrak{p}\}\,.$$

Then the following are immediate:

- $\mathcal{X} = \mathrm{Gr}_r(\mathfrak{p})$ if $r = 1$.
- $\mathcal{X} \supset \mathcal{X}_{\mathrm{ab}} := \{W \in \mathrm{Gr}_r(\mathfrak{p}) \mid W \text{ abelian}\} \simeq K/N$ where $N = N_K(\mathfrak{a})$.

If $r \geq 2$ and $\mathfrak{g}$ simple, then $\mathcal{X} \subsetneq \mathrm{Gr}_r(\mathfrak{p})$. The problem we pose is to determine $\mathcal{X}$ in general.

In this paper we describe the set $\mathcal{X}$ for $\mathfrak{g} = \mathfrak{sl}(3, \mathbb{R})$, in which case $r = 2$ and $\dim \mathrm{Gr}_2(\mathfrak{p}) = 6$. It turns out that $\mathcal{X}$ is dominated by a real algebraic variety of dimension 5: there exists a surjective algebraic map:

$$\Phi : K \times_{N_0} \mathbb{P}(\mathbb{R}^3) \to \mathcal{X}$$

with $N_0 \simeq (\mathbb{Z}/4\mathbb{Z}) \rtimes \mathbb{Z}/2\mathbb{Z}$ and generically trivial fibers (see Theorem 3.2 in Section 3 below).

In Section 4 we give an alternative approach to the problem of characterizing $\mathcal{X}$ via tools from algebraic geometry, in particular Galois-cohomology. This section evolved out of several discussions with Günter Harder and we thank him for explaining us some mathematics which was unfamiliar to us.

For general $\mathfrak{g}$ we do not know the nature of $\mathcal{X}$.

2. Description by invariants

Let $\mathfrak{g} = \mathfrak{sl}(3, \mathbb{R})$ and $\mathfrak{a} = \mathrm{diag}(3, \mathbb{R})_{\mathrm{tr}=0}$. We give the following description of $\mathcal{X}$, which will lead to the classification in the following sections.

Theorem 2.1. *The two dimensional subspace $W \in \mathrm{Gr}_2(\mathfrak{p})$ belongs to $\mathcal{X}$ if and only if it contains a non-zero matrix X with two equal eigenvalues.*

For example, with the notation in Theorem 1.1, the matrix $X_{\mu\lambda}$ with $\mu = -1$ and $\lambda = \sqrt{2}$ has eigenvalues $1, 1, -2$. Hence the space W in this theorem belongs to $\mathcal{X}$.

Proof. That this is a necessary condition is clear, since $W \in \mathcal{X}$ means that W contains at least one element from *every* K-orbit on $\mathfrak{p}$.

In order to describe the K-orbits, we recall some basic invariant theory. Let

$$u_1(X) = \mathrm{tr}\, X^2$$
$$u_2(X) = \det X$$

for $X \in \mathfrak{p}$. Then $u_1, u_2 \in \mathbb{C}[\mathfrak{p}]^K$, the ring of K-invariant polynomials on $\mathfrak{p}$. In fact, it is a well-known fact that

$$\mathbb{C}[\mathfrak{p}]^K = \mathbb{C}[u_1, u_2],$$

but we shall not use this here.

Lemma 2.2. *The level sets for $u = (u_1, u_2)$ are single K-orbits.*

Proof. Each K-orbit is uniquely determined by a set of eigenvalues (with multiplicities). It is easily seen that the characteristic polynomial of a trace free 3×3-matrix X is

$$-x^3 + \frac{1}{2} u_1(X) x + u_2(X).$$

The lemma follows immediately. □

It follows that W belongs to $\mathcal{X}$ if and only if it has a non-trivial intersection with each level set. Notice that $u_1(X)$ is the square of the trace norm of X, for X symmetric. In particular, $u_1(X) > 0$ for $X \neq 0$. Since u_1 and u_2 are homogeneous, it suffices to consider level sets of the form $\{u_1 = 1, u_2 = c_2\}$.

We thus consider for each $W \in \mathrm{Gr}_2(\mathfrak{p})$ the unit sphere

$$W_1 = \{X \in W \mid u_1(X) = 1\},$$

and we define

$$J := \{u_2(X) \mid X \in W_1\}.$$

Since W_1 is connected, J is an interval. Moreover, it is symmetric around 0, since u_2 has odd degree. In particular, we denote by

$$I := \{u_2(X) \mid X \in \mathfrak{a}_1\}$$

the interval corresponding to the unit sphere in $\mathfrak{a}$. We now show:

Lemma 2.3. *The interval I is given by $I = [-c, c]$, where $c = 54^{-1/2}$. Furthermore, the extreme values $\pm c$ are obtained precisely in those elements $X \in \mathfrak{a}_1$, which have two equal eigenvalues.*

Proof. Let us introduce coordinates for $\mathfrak{a}$, namely

$$\mathfrak{a} = \{D_{xy} := \mathrm{diag}(x, y, -x-y) \mid x, y \in \mathbb{R}\}.$$

Furthermore, let us introduce two functions:

$$f_1(x, y) := u_1(D_{xy}) = 2(x^2 + y^2 + xy)$$
$$f_2(x, y) := u_2(D_{xy}) = -xy(x + y).$$

We wish to maximize/minimize f_2 under the condition of $f_1 = 1$. For that we perform the method of Lagrange: $df_1 = 2(2x + y, 2y + x)$ and $df_2 = -(y(2x+y), x(2y+x))$ have to be collinear. This can only happen in three cases: $x = y$, $2x + y = 0$ or $2y + x = 0$. Notice that these are exactly the cases in which two of the diagonal entries of D_{xy} are equal.

We start with $x = y$. Here $f_1(x, x) = 6x^2 = 1$ means that $x = \pm 6^{-1/2}$. Hence $f_2(x, x) = -2x^3 = \pm 54^{-1/2}$. Secondly, if $2x + y = 0$,

then $f_1(x, -2x) = 6x^2 = 1$, so again $x = \pm 6^{-1/2}$. Hence $f_2(x, -2x) = -2x^3 = \pm 54^{-1/2}$. Finally, the case $2y + x = 0$ is similar. $\square$

In order to complete the proof of Theorem 2.1, we only have to note that, as $\mathfrak{p} = K \cdot \mathfrak{a}$ we have $J \subset I$ and equality $J = I$ holds if and only if $W \in \mathcal{X}$. $\square$

Remark 2.4. *(a) It follows from Theorem 2.1 that not all 2-dimensional subspaces $W \subset \mathfrak{p}$ belong to $\mathcal{X}$. An extreme case is*

$$W = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & \mu \\ 0 & \mu & -\lambda \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\} .$$

for which $u_2(W) = \{0\}$ and hence $W \notin \mathcal{X}$.
(b) Let us define a continuous function on $\mathrm{Gr}_2(\mathfrak{p})$ *by*

$$f : \mathrm{Gr}_2(\mathfrak{p}) \to \mathbb{R}_{\geq 0}, \quad W \mapsto \max_{X \in W_1} u_2(X)$$

Then we get

$$\mathcal{X} = \{W \in \mathrm{Gr}_2(\mathfrak{p}) \mid f(W) = 54^{-1/2}\}$$

by our previous result. In particular, $\mathcal{X}$ is a closed subset of $\mathrm{Gr}_2(\mathfrak{p})$.

3. K-orbits on $\mathcal{X}$

We aim to describe $\mathcal{X}$ explicitly. Our starting point is the following observation. Let

$$X_0 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} ,$$

and observe that the two dimensional subspace $W \in \mathrm{Gr}_2(\mathfrak{p})$ belongs to $\mathcal{X}$ if and only if it contains a vector in the K-orbit of X_0.

In fact this is an immediate consequence of our discussion in the previous section. Since we consider trace free matrices, X has two equal eigenvalues if and only if its eigenvalues are $\nu, \nu, -2\nu$ for some $\nu \in \mathbb{R}$, that is, X is conjugate to νX_0.

Let Ω denote the 3-dimensional variety

$$\Omega = \{W \in \mathrm{Gr}_2(\mathfrak{p}) \mid X_0 \in W\},$$

in $\mathrm{Gr}_2(\mathfrak{p})$. The stabilizer $H \subset K$ of the line $\mathbb{R}X_0$ acts on Ω. We have proved the following result:

Proposition 3.1. *The map $(k, W) \mapsto k \cdot W$ from $K \times_H \Omega$ to $\mathcal{X}$ is surjective.*

Notice that $H = MT$ where T is the maximal torus

$$T := \begin{pmatrix} \mathrm{SO}(2,\mathbb{R}) & 0 \\ 0 & 1 \end{pmatrix}$$

and $M \simeq [\mathbb{Z}/2Z]^2$ is the diagonal group group generated by

$$m_1 := \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad m_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

in K.

For $Y \notin \mathbb{R}X_0$ we set

$$W_Y := \mathrm{span}_{\mathbb{R}}\{X_0, Y\},$$

then $\Omega = \{W_Y \mid Y \notin \mathbb{R}X_0\}$. Since X_0 is fixed under the maximal torus T, we have that W_Y and $W_{t\cdot Y}$ belong to the same K-orbit for $t \in T$. Thus it suffices to consider elements Y of the following shape:

$$Y = Y_{\alpha,\delta,\epsilon} = \begin{pmatrix} \alpha & 0 & \delta \\ 0 & -\alpha & \epsilon \\ \delta & \epsilon & 0 \end{pmatrix}.$$

Let $\mathcal{Y}$ denote the 2-dimensional projective space of these lines and consider the algebraic mapping

$$K \times \mathcal{Y} \to \mathcal{X}, \quad (k, [Y]) \mapsto k \cdot W_Y . \tag{3.1}$$

The group H does not act on $\mathcal{Y}$. However, let N_0 denote the subgroup of order 8, generated by

$$s_0 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

and m_2 (note that $m_1 = s_0^2$). It follows from the relations

$$s_0 \cdot Y_{\alpha,\delta,\epsilon} = Y_{-\alpha,-\epsilon,\delta}, \quad m_2 \cdot Y_{\alpha,\delta,\epsilon} = Y_{\alpha,-\delta,\epsilon}. \tag{3.2}$$

that N_0 acts on $\mathcal{Y}$. Conversely, if $k \in H$ and $k \cdot Y \in \mathcal{Y}$ for some $Y = Y_{\alpha,\delta,\epsilon} \in \mathcal{Y}$ with $\alpha \neq 0$, then $k \in N_0$.

Since furthermore $k \cdot W_Y = W_{k\cdot Y}$ for $Y \in \mathcal{Y}$ and $k \in N_0$, the above map (3.1) factorizes to an algebraic map

$$K \times_{N_0} \mathcal{Y} \to \mathcal{X}.$$

This map is K-equivariant, continuous and onto and we wish to show that it is generically injective.

We define the following open dense subset of $\mathcal{Y}$:

$$\mathcal{Y}' = \{[Y_{\alpha,\delta,\epsilon}] \mid \alpha \neq 0,\ \delta \neq 0,\ \epsilon \neq 0\} \tag{3.3}$$

and note that it is preserved by N_0. We set $\mathcal{X}' = K \cdot \mathcal{Y}' \subset \mathcal{X}$.

Theorem 3.2. *The map*

$$K \times_{N_0} \mathcal{Y}' \to \mathcal{X}', \quad [k, [Y]] \mapsto k \cdot W_Y$$

is a K-equivariant continuous bijection. In particular $\mathcal{X}'$ carries a natural structure of a smooth 5-dimensional K-manifold.

In order to obtain this, we study the intersection of the K-orbit of X_0 with W_Y. We first prove:

Lemma 3.3. *Assume*

$$X := \begin{pmatrix} \lambda & 0 & \delta \\ 0 & \mu & \epsilon \\ \delta & \epsilon & -(\lambda+\mu) \end{pmatrix} \in K \cdot X_0.$$

Then $\delta = 0$ or $\epsilon = 0$.

Proof. It follows from Lemma 2.2 that

$$u_1(X) = 2(\lambda^2 + \mu^2 + \lambda\mu + \delta^2 + \epsilon^2) = u_1(X_0) = 6 \tag{3.4}$$

$$u_2(X) = -\lambda\mu(\lambda+\mu) - \epsilon^2\lambda - \delta^2\mu = u_2(X_0) = -2\,. \tag{3.5}$$

In particular, it follows from (3.4) that $\lambda^2 + \mu^2 + \lambda\mu \leq 3$. Since $\lambda^2 + \mu^2 + \lambda\mu = (\lambda + \frac{1}{2}\mu)^2 + \frac{3}{4}\mu^2$ this implies $|\mu| \leq 2$.

Multiplying by $\frac{1}{2}\mu$ in (3.4) and adding (3.5) we obtain

$$\mu^3 + \epsilon^2\mu - \epsilon^2\lambda = 3\mu - 2\,,$$

or equivalently

$$\epsilon^2(\lambda - \mu) = (\mu+2)(\mu-1)^2\,. \tag{3.6}$$

In particular it follows that $\epsilon = 0$ or $\lambda \geq \mu$.

Since μ and λ appear symmetrically in (3.4) and (3.5) we obtain similarly $|\lambda| \leq 2$,

$$\delta^2(\mu - \lambda) = (\lambda+2)(\lambda-1)^2\,, \tag{3.7}$$

and conclude that $\delta = 0$ or $\mu \geq \lambda$.

Notice finally that if $\lambda = \mu$, then $\lambda = \mu = -2$ or $\lambda = \mu = 1$ by (3.6), and from (3.4) it then follows that $\lambda = \mu = 1$ and $\delta = \epsilon = 0$. $\square$

We can now prove Theorem 3.2.

Proof. It remains to be seen that $k \cdot W_Y = W_{Y'}$ implies $k \in N_0$ and $[Y'] = [k \cdot Y]$ for $Y, Y' \in \mathcal{Y}'$. In particular, it follows from $k \cdot W_Y = W_{Y'}$ that $k \cdot X_0 = aX_0 + bY'$ for some $a, b \in \mathbb{R}$. Since $Y' \in \mathcal{Y}'$ it follows from Lemma 3.3 that $b = 0$, and hence $k \in H$. Now $k \cdot Y$ must be a

multiple of Y', by orthogonality with X_0 with respect to the trace form $\langle A, B\rangle = \mathrm{tr}(AB)$. It follows that $k \in N_0$. □

In order to give a complete classification of $\mathcal{X}$, one needs to describe the fibers in $K \times \mathcal{Y}$ above the elements outside of $\mathcal{X}'$. We omit the details, but mention that in general the fibers will not be finite.

4. Alternative approach

In the following two subsections We describe an alternative approach to the elements in $\mathcal{X}$, based on results from algebraic geometry and possibly useful in the general case.

4.1. Generic subspaces

Let $L \in \mathrm{Gr}_2(\mathfrak{p})$. The following property of L is closely related to the property that $K \cdot L = \mathfrak{p}$. We say that L is *generic* if there exists an element $Z \in L$ such that

$$[\mathfrak{k}, Z] + L = \mathfrak{p}. \tag{4.1}$$

By reason of dimension, the sum is necessarily direct if (4.1) holds. Equivalent with (4.1) is that the map $(k, W) \mapsto k \cdot W$ is submersive at $(1, Z)$. It follows that L is generic if and only if the image $K \cdot L$ has non-empty interior in $\mathfrak{p}$. In particular, if $L \in \mathcal{X}$, then L is generic.

It follows from Proposition 3.1 that $W_Y := \mathrm{span}_{\mathbb{R}}\{X_0, Y\}$ is generic for every $Y \in \mathcal{Y}$. As we want to proceed independently of the computations in Section 2, we sketch a simple proof of this fact. Let $Y = Y_{\alpha,\delta,\epsilon}$ where $(\alpha, \delta, \epsilon) \neq (0, 0, 0)$, and put $Z = Y + cX_0$ where $c \in \mathbb{R}$. We claim that (4.1) holds for some c.

Let $\mathfrak{k} = \mathrm{span}_{\mathbb{R}}(X_1, X_2, X_3)$ where

$$X_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad X_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Then

$$[X_1, Z] = \begin{pmatrix} 0 & -2\alpha & \epsilon \\ -2\alpha & 0 & -\delta \\ \epsilon & -\delta & 0 \end{pmatrix}$$

$$[X_2, Z] = \begin{pmatrix} 2\delta & \epsilon & -3c-\alpha \\ \epsilon & 0 & 0 \\ -3c-\alpha & 0 & -2\delta \end{pmatrix}$$

$$[X_3, Z] = \begin{pmatrix} 0 & \delta & 0 \\ \delta & 2\epsilon & -3c+\alpha \\ 0 & -3c+\alpha & -2\epsilon \end{pmatrix}$$

Hence the condition that $[X_1, Z]$, $[X_2, Z]$, $[X_3, Z]$, X_0 and Y are linearly independent amounts to

$$\det \begin{pmatrix} 0 & -2\alpha & \epsilon & 0 & -\delta \\ 2\delta & \epsilon & -3c-\alpha & 0 & 0 \\ 0 & \delta & 0 & 2\epsilon & -3c+\alpha \\ 1 & 0 & 0 & 1 & 0 \\ \alpha & 0 & \delta & -\alpha & \epsilon \end{pmatrix} \neq 0.$$

This determinant is a second order polynomial in c. It is easily seen that the coefficient of c^2 is $18\alpha^2$. On the other hand, if $\alpha = 0$ then the constant term in the polynomial is $2(\delta^2 + \epsilon^2)^2$. In any case, it is a non-zero polynomial in c, and our claim is proved.

It is of interest also to see which other spaces L are generic.

Lemma 4.1. *Let $L \in \mathrm{Gr}_2(\mathfrak{p})$. Then L is generic if and only if it is conjugate to* $\mathrm{span}_{\mathbb{R}}\{X, Y\}$*, where*

$$X := \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & -\lambda-\mu \end{pmatrix}, \quad Y := \begin{pmatrix} \alpha & \gamma & \delta \\ \gamma & \beta & \epsilon \\ \delta & \epsilon & -\alpha-\beta \end{pmatrix}$$

and either

(i) *two of the elements $\lambda, \mu, -\lambda-\mu$ in X are equal*

or

(ii) $\alpha\mu - \beta\lambda \neq 0$.

Proof. Before the proof we make the following observation. Let $Z \in L$ where $L \in \mathrm{Gr}_2(\mathfrak{p})$ is arbitrary. It follows easily from the relation $\mathrm{tr}([U, Z]V) = \mathrm{tr}(U[Z, V])$ for $U \in \mathfrak{k}$ and $Z, V \in \mathfrak{p}$, that $[\mathfrak{k}, Z]$ can be characterized as the set of elements $T \in \mathfrak{p}$, for which $\mathrm{tr}(TV) = 0$ for all V in the centralizer of Z in $\mathfrak{p}$.

Assume now that $L = \mathrm{span}_{\mathbb{R}}\{X, Y\}$ as above. In case (i), L belongs to $\mathcal{X}$ and is generic as established above. Assume (ii) and not (i). Since

the diagonal elements of X are mutually different it follows that the centralizer of X in $\mathfrak{p}$ is $\mathfrak{a}$, and hence $[\mathfrak{k}, X]$ consists of the matrices in $\mathfrak{p}$ with zero diagonal entries. Since $\alpha\mu - \beta\lambda \neq 0$ it follows that X and Y are linearly independent from $[\mathfrak{k}, L]$. Hence L is generic.

Conversely, if $L \in \mathrm{Gr}_2(\mathfrak{p})$ is generic, then by conjugation we can arrange that the matrix Z in (4.1) is diagonal. Let $X = Z$. As before it follows that if (i) does not hold, then $[\mathfrak{k}, Z]$ consists of all the matrices which are zero on the diagonal. Hence any $Y \in L$ linearly independent from Z must have the mentioned form. □

For example, the subspace W in Remark 2.4 (a) is not generic. On the other hand, it is not difficult to find examples of subspaces which are generic, but do not belong to $\mathcal{X}$. For example when $\lambda = 0$, $\mu = 1$, $\alpha = 1$, $\beta = 0$, $\gamma = \delta = 0$ and $|\epsilon| > 3/2$ in the expressions above, then $L = \mathrm{span}_{\mathbb{R}}\{X, Y\}$ is generic and not in $\mathcal{X}$.

4.2. **Approach via algebraic geometry**

The following evolved from discussions with Günter Harder.

Let $\mathrm{Gr}_2(\mathfrak{p}_{\mathbb{C}})$ be the complex variety of 2-dimensional complex subspaces of $\mathfrak{p}_{\mathbb{C}}$. For a subspace $L \in \mathrm{Gr}_2(\mathfrak{p}_{\mathbb{C}})$ we denote by $\overline{L}$ its complex conjugate. Note that there is a natural embedding $\mathrm{Gr}_2(\mathfrak{p}) \hookrightarrow \mathrm{Gr}_2(\mathfrak{p}_{\mathbb{C}})$ the image of which,

$$\mathrm{Gr}_2(\mathfrak{p}) = \{L \in \mathrm{Gr}_2(\mathfrak{p}_{\mathbb{C}}) \mid L = \overline{L}\},$$

constitutes the real points of $\mathrm{Gr}_2(\mathfrak{p}_{\mathbb{C}})$. Let $L_{\mathbb{R}} = L \cap \mathfrak{p} \in \mathrm{Gr}_2(\mathfrak{p})$ for $L \in \mathrm{Gr}_2(\mathfrak{p}_{\mathbb{C}})$ with $\overline{L} = L$.

As before we call $L \in \mathrm{Gr}_2(\mathfrak{p}_{\mathbb{C}})$ *generic* if there exist $Z \in L$ such that

$$L + [\mathfrak{k}_{\mathbb{C}}, Z] = \mathfrak{p}_{\mathbb{C}}$$

or equivalently, such that

$$\Phi_L : K_{\mathbb{C}} \times L \to \mathfrak{p}_{\mathbb{C}}, \quad (k, z) \mapsto k \cdot z$$

is submersive at $(1, Z)$. It is clear that the complexification of a generic subspace in $\mathfrak{p}$ is generic. Let us remark that if L is generic then

- The image $\operatorname{im} \Phi_L$ has non-empty Zariski open interior in $\mathfrak{p}_{\mathbb{C}}$.
- The orbit $\mathcal{O}_L := K_{\mathbb{C}} \cdot L \in \mathrm{Gr}_2(\mathfrak{p}_{\mathbb{C}})$ satisfies:

$$\dim_{\mathbb{C}} \mathcal{O}_L = \dim K_{\mathbb{C}} = 3\,.$$

Let now L be generic. To $\mathcal{O}_L$ we associate:

$$\mathcal{Z}_L := \{(z, W) \mid z \in W,\ W \in \mathcal{O}_L\}\,.$$

The projection onto the second factor $\pi_2 : \mathcal{Z}_L \to \mathcal{O}_L$ reveals the structure of an algebraic $\mathbb{C}^2$-vector bundle over $\mathcal{O}_L$. In particular $\dim_\mathbb{C} \mathcal{Z}_L = 5$. The projection onto the first factor $\pi_1 : \mathcal{Z}_L \to \mathfrak{p}_\mathbb{C}$ features

$$\operatorname{im} \pi_1 = \operatorname{im} \Phi_L \, .$$

In particular $\operatorname{im} \pi_1$ contains a non-empty Zariski-open set.

Now let us assume that L is the complexification of a generic subspace in $\mathfrak{p}$. Then $\mathcal{O}_L$ and $\mathcal{Z}_L$ are defined over $\mathbb{R}$. The real points of $\mathcal{Z}_L$ are given by

$$\mathcal{X}_L := \mathcal{Z}_L^\mathbb{R} = \{(x, W) \in \mathcal{Z}_L \mid W = \overline{W},\ x \in W_\mathbb{R}\} \, .$$

Again $\pi_1(\mathcal{X}_L) \subset \mathfrak{p}$ is a constructible set with non-empty open interior.

In order to determine $\mathcal{X}_L$ we have to determine the real points of $\mathcal{O}_L$. In general this is a finite union of K-orbits which is difficult to determine as one needs to know the $K_\mathbb{C}$-stabilizer of L.

However, if we suppose that L is such that $\mathcal{O}_L \simeq K_\mathbb{C}$, then the real points of $\mathcal{O}_L$ are just $K \cdot L$ and $\operatorname{im} \pi_1(\mathcal{X}_L) = K \cdot L_\mathbb{R} \subset \mathfrak{p}$. Hence $K \cdot L_\mathbb{R}$ has non-empty Zariski open interior and thus $K \cdot L_\mathbb{R} = \mathfrak{p}$ as the left hand side is closed. We have thus established that $L \in \mathcal{X}$ for every generic $L \in \mathrm{Gr}_2(\mathfrak{p})$ with trivial stabilizer in $K_\mathbb{C}$.

Notice that the stabilizer in $K_\mathbb{C}$ is trivial if the stabilizer in K is trivial. This can be seen as follows. Let us denote by $S \subset K_\mathbb{C}$ the stabilizer of L. As L is generic, S is a discrete subgroup of $K_\mathbb{C}$. We have to show that S is trivial if $S \cap K$ is trivial. Note that $S = \overline{S}$. Therefore, for $k \in S$,

$$x := \overline{k}^{-1} k \in S \, .$$

Observe that $x = \exp(X)$ for a unique $X \in i\mathfrak{k}$. As x is positive definite, it follows from $x \cdot L = L$ that $\exp(\mathbb{R}X) \subset S$. Thus $X = 0$ by the discreteness of S, and hence $k = \overline{k}$.

Let $\mathcal{Y}''$ denote the following subset of $\mathcal{Y}' \subset \mathcal{Y}$

$$\mathcal{Y}'' = \{[Y_{\alpha,\delta,\epsilon}] \mid \alpha \neq 0,\ \delta \neq 0,\ \epsilon \neq 0,\ \delta \neq \pm\epsilon\} \tag{4.2}$$

We claim that for $[Y] \in \mathcal{Y}''$, the K-stabilizer of W_Y is trivial. Assume $k \cdot W_Y = W_Y$ for some $k \in K$. In the proof of Theorem 3.2 we saw that $k \in N_0$ and $k \cdot Y = \pm Y$, and then it follows from (3.2) that $k = e$.

To summarize, we have shown with alternative methods that $W_Y \in \mathcal{X}$ for $Y \in \mathcal{Y}''$.

Let us now deal with generic orbits $\mathcal{O} := \mathcal{O}_L$ where the $K_\mathbb{C}$-stabilizer is not necessarily trivial. For that we first have to recall the concept of non-abelian cohomology (see [2], Sect. 5.1).

Let Γ and H be groups. We assume that Γ acts on H by preserving the group law of H, in other words: there exists a homomorphism $\alpha : \Gamma \to \mathrm{Aut}(H)$. In the sequel we write for $g \in \Gamma$ and $h \in H$

$$^{g}h := \alpha(g)(h)\,.$$

By a cocycle of Γ in H we understand a map

$$\theta : \Gamma \to H, \quad g \mapsto \theta(g)$$

such that

$$\theta(g_1 g_2) = \theta(g_1) \cdot {}^{g_1}\theta(g_2)\,.$$

The set of all cocycles is denoted by $Z^1(\Gamma, H)$. We call two cocycles θ, θ' homologous if there exists an $h \in H$ such that

$$\theta'(g) = h^{-1}\theta(g)^{g}h$$

for all $g \in \Gamma$. The corresponding set of equivalence classes $H^1(\Gamma, H) = Z^1(\Gamma, H)/\sim$ is refered to as the first cohomology set of Γ with values in H.

Henceforth we let $\Gamma = \mathrm{Gal}(\mathbb{C}|\mathbb{R})$ be the Galois group of $\mathbb{C}|\mathbb{R}$. We write $\Gamma = \{1, \sigma\}$ with σ the non-trivial element. Note that Γ acts on $K_{\mathbb{C}}$ by complex conjugation. In fact σ induces the Cartan involution on $K_{\mathbb{C}}$. Likewise Γ acts on the stabilizer $S < K_{\mathbb{C}}$ of L.

Let us denote by $\mathcal{O}(\mathbb{R})$ the real points of $\mathcal{O}$ and by $[\mathcal{O}(\mathbb{R})] := \mathcal{O}(\mathbb{R})/K$, the set of all K-orbits. Then for $k \in K_{\mathbb{C}}$ such that $z = k \cdot L \in \mathcal{O}(\mathbb{R})$ we define a cocycle

$$\theta_k(\sigma) := \sigma(k)^{-1}k\,.$$

Replacing k by ks for $s \in S$ results in $\theta_{ks}(\sigma) = \sigma(s)^{-1}\sigma(k)^{-1}ks$ which is homologous to θ_k. Hence the prescription $\theta_z := \theta_k$ gives us a well defined element in $H^1(\Gamma, S)$. Further note that $\theta_{kz} = \theta_z$ for all $k \in K$. Therefore we obtain a map

$$\Phi : [\mathcal{O}(\mathbb{R})] \to H^1(\Gamma, S), \quad [z] \mapsto \theta_z\,.$$

It is easy to check that Φ is injective and it remains to characterize the image. For that we consider the natural map between pointed sets

$$\Psi : H^1(\Gamma, S) \to H^1(\Gamma, K_{\mathbb{C}})\,.$$

Define $\ker\Psi := \Psi^{-1}(\mathbf{1})$. We recall that twisting (cf. [2], Sect. 5.4) implies that all pre-images of Ψ are in fact kernels with S and $K_{\mathbb{C}}$ replaced by twists.

We claim that $\mathrm{Im}\,\Phi = \ker\Psi$. The inclusion "$\subset$" is obvious. Suppose that $\Psi(\theta) = \mathbf{1}$ is the trivial cocycle. Then $\theta(\sigma) = \sigma(k)^{-1}k$ for some $k \in K_{\mathbb{C}}$. As $\theta(\sigma) \in S$ it follows that $k \cdot L \in \mathcal{O}(\mathbb{R})$ and our claim is established.

We have thus shown that

$$[\mathcal{O}(\mathbb{R})] \simeq \ker\big(H^1(\Gamma, S) \to H^1(\Gamma, K_{\mathbb{C}})\big) \tag{4.3}$$

and thus

$$K \cdot L = \mathfrak{p} \quad \iff \quad \ker\big(H^1(\Gamma, S) \to H^1(\Gamma, K_{\mathbb{C}})\big) = \{\mathbf{1}\}\,. \tag{4.4}$$

In the next step we wish to characterize the cohomology sets involved. For $H^1(\Gamma, K_{\mathbb{C}})$ we can use Harder's Theorem (cf. [1], Theorem III) to obtain

$$H^1(\Gamma, K_{\mathbb{C}}) = \mathbb{Z}/2\mathbb{Z}\,. \tag{4.5}$$

In order to discuss the structure of $H^1(\Gamma, S)$ we mention its more convenient description as a subset of S:

$$H^1(\Gamma, S) = \{s \in S \mid s\sigma(s) = 1\}/\sim$$

where $s \sim s'$ if $s' = hs\sigma(h)^{-1}$ for some $h \in S$. Now the fact that L is generic and S is σ-stable implies that $S \subset K$ (see our argument from above). Hence

$$H^1(\Gamma, S) = \{s \in S \mid s^2 = 1\}/\sim$$

with $s \sim s'$ if $s' = hsh^{-1}$ for some $h \in S$.

To see an example let us consider L's which correspond to subspaces $\mathcal{Y}'\backslash\mathcal{Y}''$. For those L the stabilizer is contained in $N_0 \simeq (\mathbb{Z}/4\mathbb{Z})\rtimes(\mathbb{Z}/2Z)$. Actually the stabilizer is either $\{\mathbf{1}, m_2s_0\}$ or $\{\mathbf{1}, s_0m_2\}$. Let us consider the first case: with $\gamma := m_2s_0$

$$\gamma = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

We have to show that $\gamma \notin \ker\Psi$. Now $\gamma \in \ker\Psi$ means $\gamma = k\sigma(k)^{-1}$ for some $k \in K_{\mathbb{C}}$, or equivalently

$$\gamma\sigma(k) = k\,.$$

This means that the last row of k consists of imaginary elements; a contradiction to the fact that the sum of their squares adds up to one. Hence $\ker\Psi$ is trivial and therefore $K \cdot L = \mathfrak{p}$.

References

[1] T. Geisser, *Galoiskohomologie reeller halbeinfacher algebraischer Gruppen*, Abh. Math. Sem. Univ. Hamburg **61** (1991), 231–242

[2] J.-P. Serre, *Galois Cohomology*, Springer Monographs in Math., 1996

Max-Planck-Institut für Mathematik, Vivatsgasse 7, D-53111 Bonn, Germany, email: kroetz@mpim-bonn.mpg.de

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 København, Denmark, email: schlicht@math.ku.dk